Mannerfelt et al. use 21,703 historical terrestrial photographs, acquired during the period 1916-1947, to reconstruct the 3D geometry of 45% of Switzerland’s glaciated area nearly a century ago. This paper is an absolute tour de force. I was impressed by not only the sheer scale of the analysis (successfully generating 113 different Agisoft Metashape models, each with an average of 192 historical photographs), but also the level of care and detail in the error propagation and uncertainty analysis. This project provides an impressive and carefully curated dataset that will prove useful to the glaciology community for better understanding the drivers of glacier mass loss. I have two general suggestions, followed by a handful of more specific comments, that I hope can help Mannerfelt et al. improve the presentation of this exciting manuscript. First, I encourage Mannerfelt et al. to dedicate a paragraph or two in the discussion to explore the implications of their new mass balance constraints. What have we learned from this new dataset? For example, do the authors have ideas about what controls the regional mass balance patterns shown in Fig. 7? What fraction of that mass balance variability can be explained by temperature/elevation, precipitation, glacier slope, debris cover, etc.? How might these climatic and geometric controls on glacier mass balance strengthen or weaken in the next century? The authors don’t need to address all of these questions, but I encourage them to take stock of their hard-earned new dataset and explore some of the implications. Second, I thought the description of the estimation and propagation of different sources of uncertainty (section 3) was excellent. It was concise and efficient, yet detailed enough to be replicated. Well done. I only have one suggestion, which is that I think the elevation change rate (m/yr) uncertainty for *each individual glacier* should be better conveyed to the reader. For example, Fig. 4 should have another panel showing a histogram of the cumulative uncertainty (all sources) for each glacier. Then, the median of this distribution can be reported in the abstract, results, etc. For example, the Swiss-wide mean glacier mass balance was -0.52 +/- 0.09 m/yr during the period 1931-2016. On average, for an individual glacier, what is the uncertainty in glacier mass balance? Is it 0.1 m/yr, 0.6 m/yr, etc.? I think this is an important number to convey clearly, since it helps
the reader understand how much uncertainty is associated with an individual glacier mass balance estimate (rather than the regional and Swiss-wide averages). Overall, this is a very impressive and exciting manuscript, and I recommend publication with minor revisions.

Specific comments

Lines 21-22: It might provide useful context for the reader if you plot the glaciological mass balance data for the handful of monitored glaciers in Switzerland and the European Alps—for example, can Fig. 1 include a sub-panel that has a plot like Fig. 5 of Huss et al. (2015) or Fig. 1 in Oerlemans (1994), with the published of field-based glacier change data for European glaciers? I think that plot would effectively convey both how sparse the existing data are (i.e., how few glaciers have repeat field-based monitoring), and a sense of the patterns and magnitude of glacier change over the last century.

Lines 76-77: “The cameras were of two different brands called “Wild” and “Zeiss”, which had different focal lengths, image dimensions, and frame appearances…” Perhaps you should be more specific here—you explain that the cameras are of two different brands, but are there multiple camera types and lenses of each brand? In other words, do you solve for a full separate set of camera parameters for each of the 21 individual cameras used, or, are there camera distortion parameters that are largely shared between the two groups of Wild vs. Zeiss cameras? You mostly answer this question in lines 165-170, but I’m still curious whether you shared any of the camera parameters across different individual cameras of the same model. Alternatively, after the fact, did you explore how different the Metashape-determined camera parameters varied across individual cameras of the same model? That would be an interesting observation to share with folks who want to do similar regional reconstructions for other glaciated areas.

Line 116: “In a nutshell, images and associated metadata are preprocessed first for homogenizing the inputs.” – Can you be a little more specific here? What are you homogenizing? Perhaps you should say “…images are preprocessed first to remove the geometric distortions introduced during digitization (scanning) and to correct for biased position data.”

Lines 135-137: Can you give the reader a little bit of intuition for why, in 32% of the images in the dataset, only 3 fiducial marks could be identified? Were the images too degraded to see the pattern by eye, or was it just that the automated detection couldn’t spot the fiducial marks in regions of the images that were very dark, etc.?

Lines 144-149: Can you add a sentence to explain how the initial position data from ~1931 was determined? Why does this slight bias exist in the position data?
Lines 167-170: I’m interested and a little surprised that the same individual camera can have different parameters (as constrained during the Metashape reconstruction) from different years! How much do the distortion parameters vary? Is it more than you expect from the uncertainties associated with the Metashape reconstructions? Have you done any sensitivity experiments to make sure that you’re not overfitting the camera parameters in Metashape due to the inherent limitations of the dataset—minimal spatial coverage, physical image degradation, poor overlap, etc.?

Lines 208-210: Can you give the reader a little intuition for the reason of the limited glacier coverage (55% of the glaciated area is missing). Is this largely because much of the glacier area was not imaged from 2 viewing angles (due to occlusion and a limited number of photograph stations?), or, is it that much of the glacier surface didn’t have enough “features” (texture) that could be matched during the photogrammetric reconstruction? If it’s the latter, it seems surprising that the upper-elevation areas, with more snow/firn cover, wouldn’t have sparser coverage?

Lines 321-322: Can you put this number in context? What percent of the total eustatic sea level rise (1931-2016) is represented by the melting of Swiss glaciers?

Lines 331-334: Perhaps you can address this in the discussion rather than the results, but I think you should add a sentence or two to explain what you think the mechanism is to explain why ice at higher elevation is experiencing less negative mass balance, and whether you think this trend is going to be strengthened or weakened in the future.

Lines 409-410: “The uncertainty analyses of this study show that the regional aggregates are accurate, thus indicating that a reanalysis of the TerrA imagery with modern methods may be worthwhile especially if interested in individual glaciers.” See my second general comment in the first paragraph of this review. I think it would help the reader to convey more clearly what the uncertainty is for the mass balance estimate of any individual glacier. This could be conveyed effectively with a histogram where the x-axis is uncertainty (m/yr) and the y-axis is the number of glaciers (like Fig. 4C, but showing cumulative uncertainties).

Figure 1: Perhaps you should include a small inset in Fig. 1A to show the map of all European glaciers and country borders, with a rectangle highlighting your study area shown in (A)?

Figure 4: I think the most useful error-related figure for the reader would be a modification of (C) in which you plot the histogram of the *total* elevation change rate uncertainty (propagated from all the different components of uncertainty shown as separate colors in Fig. 4C) for the glaciers you reconstructed. What’s the median glacier-specific dH/dt uncertainty? That would be a really helpful statistic to report widely in the paper (e.g., in the Fig. 4 caption, in the results, and in the abstract and conclusion).
Technical comments/corrections (line edits)

Line 40: remove “that is”

Line 45: “of nearly all glaciers in Switzerland” – can you finish this sentence with the number of reconstructed glaciers in parentheses? (i.e. “...of nearly all glaciers in Switzerland (XXX)”).

Line 115: remove “being”

There’s a mix of British- vs. American-style spelling of various words—I’d stick to one or the other for consistency. It seems that the American style is used most of the time. Here are a few occurrences of British English that I saw:

Line 46: centre -> center

Line 100: centre -> center

Line 125: centrepoint -> centerpoint

Line 193: centre -> center

Line 327: analyse -> analyze

Line 331: Analysing -> Analyzing

Line 401: analysed -> analyzed

Line 129: replace “increase” with “introduce” (since 2 fiducial marks represents zero redundancy for the similarity transformation, so 3 fiducial marks introduces redundancy).
Line 342: replace “and based” with “or based.”

Line 396: remove “sheer”

Line 396: replace “Whilst” with “Although”

Line 405: “improving” -> “improve”

Line 406: “pave the way to an increased coverage when compared to...” -> “pave the way to increased coverage compared to...”

Line 411: “or as alluded to above” -> This phrase is super vague—an easy fix would be to use parentheses to cite a handful of other examples of using historical imagery to reconstruct glacier change.

Line 419: should “database” be “databases”? I think this word should be plural.

Line 432: add commas after “approach” and “ago”

References
